PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/16, C07K 14/64, C12N 5/10, 1/21, 1/19, C07K 16/26, A61K 38/22, 39/395, C07K 19/00

A2

(11) International Publication Number: WO 97/16549

(43) International Publication Date: 9 May 1997 (09.05.97)

(21) International Application Number: PCT/US96/17342

(22) International Filing Date: 1 November 1996 (01.11.96)

(30) Priority Data:

60/006,221	3 November 1995 (03.11.95)	US
60/012,016	21 February 1996 (21.02.96)	US

(71) Applicant *(for all designated States except US)*: REGENERON PHARMACEUTICALS, INC. [US/US]; 777 Old Saw Mill River Road, Tarrytown, NY 10591-6707 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: DAVIS, Samuel [US/US]; Apartment #B2, 332 W. 88th Street, New York, NY 10024 (US).

(74) Agents: COBERT, Robert, J.; Regeneron Pharmaceuticals, Inc., 777 Old Saw Mill River Road, Tarrytown, NY 10591 (US) et al.

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MOLECULAR CLONING AND CHARACTERIZATION OF MOLECULES RELATED TO RELAXIN AND THE INSULIN FAMILY OF LIGANDS

(57) Abstract

This invention concerns the molecular cloning and characterization of nucleic acids encoding molecules, referred to as relaxin-related factors, that are related to the hormone relaxin and to the insulin family of ligands. More specifically, this invention provides an isolated nucleic acid molecule encoding relaxin-related factor-1. The invention also provides an isolated nucleic acid molecule encoding relaxin-related factor-2. The present invention also provides a nucleotide sequence that encodes a relaxin-related factor, as well as cells which are genetically engineered to produce the relaxin-related factor. The invention further provides for an isolated relaxin-related factor substantially free of other proteins, as well as for an antibody which specifically binds the relaxin-related factor. The invention also provides for pharmaceutical compositions comprising a relaxin-related factor, or an antibody which specifically binds the relaxin-related factor, and a pharmaceutically acceptable carrier, as well as for methods of treatment of a human or animal body. The invention further provides for a relaxin-related factor ligandbody which comprises a relaxin-related factor fused to an immunoglobulin constant region. The invention further provides for a method of treatment comprising administering a relaxin-related factor ligandbody to a human or animal body.

A4 - 09/617,389

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Ausuralia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan

MOLECULAR CLONING AND CHARACTERIZATION OF MOLECULES RELATED TO RELAXIN AND THE INSULIN FAMILY OF LIGANDS

This application claims the priority of U.S. provisional application 60/012,016 filed February 21, 1996 and the priority of U.S. provisional application 60/006,221 filed November 3, 1995, the contents of each of which in its entirety is incorporated by reference into this application. Throughout this application various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application.

This invention concerns the molecular cloning and characterization of nucleic acid sequences that encode molecules, hereinafter referred to as relaxin-related factors, that are related to the hormone relaxin and to the insulin family of ligands. More specifically, this invention provides isolated nucleic acid molecules encoding relaxin-related factors. The nucleic acid molecules may be expressed in suitable host cells and are thus useful for the production of the relaxin-related factors and relaxin-related factor ligandbodies, as well as for the production of probes that may be used to further screen for additional related family members. The invention further provides isolated relaxin-related factors substantially free of other proteins. The isolated relaxin-related factor proteins, as well as derivatives thereof, may be used for the preparation of antibodies and therapeutic compositions and as growth factors for maintaining cultures of cells which have receptors for relaxin-related factor.

BACKGROUND OF THE INVENTION

Relaxin is a polypeptide hormone that has been found in the female of all species studied. Relaxin is synthesized and stored in the corpora lutea of ovaries during pregnancy and is released into the blood stream prior to parturition. The corpus luteum of pregnancy is the main source of relaxin in many species but in others the decidua is apparently of greater importance. Hisaw suggested an important role for relaxin in mammals through its effects in dilating the pubic symphysis, thus facilitating the birth process. (Hisaw, F. L., Proc. Soc. Exp. Biol. Med. 23, 661-

663 (1926)). Relaxin has since been found to have many other possible roles, including preparation of the endometrium for implantation, inhibition of uterine activity in early pregnancy, remodeling of the uterine stroma during pregnancy, cervical ripening and the initiation of parturition. Relaxin's main cellular action in pregnancy may be to drive collagen biosynthesis in its target organs, thus facilitating the remodeling of the connective tissue. Published PCT application WO8907945 published September 8, 1989 filed by Genentech, Inc., discloses relaxin compositions which are stated to be useful for modulating the reproductive physiology of mammals during pregnancy and parturition. Relaxin has been reported to be capable of altering the nature of connective tissue, influencing smooth muscle contraction, and to have the general properties of a growth factor. It has also been reported that relaxin may influence intraocular pressure when administered in pharmacologic doses. (Kass, M.A. et al., Survey of Opththalmology 22(3): 153-176 (1977)).

Although the corpora lutea of the ovary as well as decidual and placental tissues are the most likely sites for expression of relaxin-related genes, relaxin has also been found in other tissues such as prostatic fluid and testis. There is evidence in pigs for relaxin gene expression and translation into protein in the theca interna cells of the preovulatory follicle, the corpus luteum of the cycle and the uterus. Relaxin has been identified in boar seminal plasma and can maintain or increase sperm motility. (See Bagnell, C.A., et al., J. Reprod. & Fert. (Supp.) 48:127-138 (1993) for review). It has been reported that specific, high activity human relaxin binding sites are present in discrete regions of the rat brain such that the distribution of some of these sites may be consistent with a role for relaxin in control of vascular volume and blood pressure. (Osheroff, P.L. and Phillips, H.S., Proc. Nat. Acad. Sci. (USA) 88 (15): 6413-6417 (1991)).

Biologically active relaxin consists of two peptide chains (known as the A and B chains) held together by one intra-chain and two inter-chain disulfide bonds. The structure thus closely resembles insulin, which has led to speculation of a common ancestral gene for these hormones (Schwabe, C., et al., Biochem.

Biophys. Res. Commun. 75, 503-510 (1977); James, R., et al., Nature, 267:544-546 (1977)). Relaxin is considered a member of the insulin family, which also includes insulin, IGF1, IGF2 and Leydig insulin-like protein. cDNA clones for both rat and porcine relaxins have been obtained via recombinant DNA techniques. (Hudson, P., Haley, J., Cronk, M., Shine, J. and Niall, H. Nature, 291: 127-131 (1981)). Further, the molecular cloning and characterization of a gene sequence coding for human relaxin has been described. (U.S. Patent No. 4,871,460 issued October 3, 1989 to Hudson, et al.)

SUMMARY OF THE INVENTION

This invention concerns the molecular cloning and characterization of nucleic acids encoding molecules, referred to as relaxin-related factors, that are related to the hormone relaxin and to the insulin family of ligands.

More specifically, this invention provides an isolated nucleic acid molecule encoding relaxin-related factor-1 (RRF-1). The present invention also provides a nucleotide sequence that encodes RRF-1, as well as cells which are genetically engineered to produce RRF-1.

This invention also provides an isolated nucleic acid molecule encoding relaxin-related factor-2 (RRF-2). The present invention also provides a nucleotide sequence that encodes RRF-2, as well as cells which are genetically engineered to produce RRF-2.

The invention further provides for isolated RRF-1 or RRF-2 substantially free of other proteins, as well as for an antibody which specifically binds RRF-1 and an antibody that specifically binds RRF-2. The antibodies may be polyclonal or monoclonal.

The invention also provides for a pharmaceutical composition comprising RRF-1 or RRF-2 and a pharmaceutically acceptable carrier. The invention further

provides for a pharmaceutical composition comprising an antibody which specifically binds RRF-1 or an antibody that specifically binds RRF-2 and a pharmaceutically acceptable carrier.

The invention further provides for a method of treatment comprising administering RRF-1 or RRF-2 to an animal or human body as well as a method of treatment comprising administering to an animal or human body an antibody which specifically binds RRF-1 or RRF-2.

The invention further provides for a relaxin-related factor ligandbody which comprises RRF-1 or RRF-2 fused to an immunoglobulin constant region. The invention further provides for a method of treatment comprising administering a relaxin-related factor ligandbody to a human or animal body.

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 - Nucleotide and deduced amino acid (single letter code) sequences of relaxin-related factor-1 (RRF-1).

FIGURE 2 - Comparative alignment of amino acid sequences of relaxin-related factor-1 (RRF-1) and relaxin-related factor-2 (RRF-2) with amino acid sequences of other members of the insulin family including: relaxin H1; relaxin H2; Leydig insulin-like protein (leydig); insulin; insulin-like growth factor 1 (IGF1); and insulin-like growth factor 2 (IGF2). Conserved cysteine residues are indicated by asterisks above the RRF-1 sequence. Dots indicate identical amino acid residues. Dashes represent artificial insertions for purposes of alignment. Deletions ranging from 14 - 123 amino acids as well as deletions of the signal sequences have been made for purposes of alignment and are indicated in parentheses.

FIGURE 3 - Northern blot showing expression of message encoding relaxin-related factor-1 (RRF-1) in human testis, but not in any other human tissue tested. Lanes, from left, are Brain, Cerebellum, Thymus, Heart, Liver, Small

Intestine, Skeletal Muscle, Bone Marrow, Prostate, Testis, Ovary, Uterus, Placenta.

FIGURE 4 - Nucleotide and deduced amino acid (single letter code) sequences of relaxin-related factor-2 (RRF-2).

FIGURE 5 - Northern blot showing expression of message encoding relaxin-related factor-2 (RRF-2) in human placenta, but not in any other human sample tested. Lanes, from left, are Cerebellum, Retina, Thymus, Heart, Lung, Spleen, Liver, Kidney, Adrenal, Pancreas, Small Intestine, Testis, Skeletal Muscle, HeLa cells, Placenta, Fetal Brain and Fetal Liver.

DETAILED DESCRIPTION OF THE INVENTION

This invention concerns the molecular cloning and characterization of nucleic acids encoding molecules, hereinafter referred to as relaxin-related factors, that are related to the hormone relaxin and to the insulin family of ligands. More specifically, this invention provides an isolated nucleic acid molecule encoding relaxin-related factor-1 (RRF-1). The present invention also provides a nucleotide sequence that encodes RRF-1 as well as cells which are genetically engineered to produce the molecule, by e.g. transfection, transduction, infection, electroporation, or microinjection of a nucleic acid encoding RRF-1 as described herein in a suitable expression vector.

This invention also provides an isolated nucleic acid molecule encoding relaxin-related factor-2 (RRF-2). The present invention also provides a nucleotide sequence that encodes RRF-2 as well as cells which are genetically engineered to produce the molecule, by e.g. transfection, transduction, infection, electroporation, or microinjection of a nucleic acid encoding RRF-2 as described herein in a suitable expression vector.

The invention further provides for an isolated nucleic acid molecule encoding relaxin-related factor-1, having a sequence selected from the group consisting of:

(a) the nucleotide sequence comprising the coding region of relaxin-related factor-1 as set forth in Figure 1;

(b) a nucleotide sequence that hybridizes under moderately stringent conditions to the nucleotide sequence of (a) and which encodes a molecule having the biological activity of relaxin-related factor-1; and

(c) a nucleotide sequence which, but for the degeneracy of the genetic code would hybridize to a nucleotide sequence of (a) or (b), and which encodes a molecule having the biological activity of relaxin-related factor-1.

The invention further provides for an isolated nucleic acid molecule encoding relaxin-related factor-2, having a sequence selected from the group consisting of:

(a) the nucleotide sequence comprising the coding region of relaxin-related factor-2 as set forth in Figure 4;

(b) a nucleotide sequence that hybridizes under moderately stringent conditions to the nucleotide sequence of (a) and which encodes a molecule having the biological activity of relaxin-related factor-2; and

(c) a nucleotide sequence which, but for the degeneracy of the genetic code would hybridize to a nucleotide sequence of (a) or (b), and which encodes a molecule having the biological activity of relaxin-related factor-2.

The invention also provides for a vector which comprises an isolated nucleic acid molecule encoding relaxin-related factor-1. The invention further provides for a vector which comprises an isolated nucleic acid molecule encoding relaxin-related factor-2. In one embodiment, the vector is operatively linked to an expression control sequence capable of directing its expression in a host cell. In another embodiment, the vector is a plasmid.

The invention further provides for a host-vector system for the production of a relaxin-related factor, polypeptide, or peptide fragment thereof which comprises a vector of the invention in a host cell. The host cell preferably may be a bacterial cell such as E. coli, or a yeast, insect or mammalian cell. The invention further provides for a method of producing a relaxin-related factor which comprises

growing cells of a host-vector system under conditions permitting production of the relaxin-related factor, and recovering the relaxin-related factor so produced. In one embodiment of the invention the relaxin-related factor is RRF-1. In another embodiment, the relaxin-related factor is RRF-2.

The invention further provides for an isolated relaxin-related factor substantially free of other proteins. In one embodiment, the isolated relaxin-related factor is RRF-1 and is encoded by the nucleotide sequence comprising the coding region of RRF-1 as set forth in Figure 1. In another embodiment, the isolated relaxin-related factor is RRF-2 and is encoded by the nucleotide sequence comprising the coding region of RRF-2 as set forth in Figure 4.

The invention further provides for an antibody which specifically binds a relaxin-related factor. More specifically, the invention provides for an antibody which specifically binds RRF-1 and an antibody that specifically binds RRF-2. The antibody may be a polyclonal or a monoclonal antibody.

The invention further provides for a pharmaceutical composition comprising a relaxin-related factor and a pharmaceutically acceptable carrier. By way of non-limiting example, the relaxin-related factor may be RRF-1 or RRF-2. The invention further provides for a pharmaceutical composition comprising an antibody which specifically binds a relaxin-related factor and a pharmaceutically acceptable carrier. The relaxin-related factor that the antibody specifically binds may be RRF-1 or RRF-2.

The invention further provides for a method of treatment comprising administering a relaxin-related factor to an animal or human body. The invention also provides for a method of treatment comprising administering an antibody which specifically binds a relaxin-related factor to an animal or human body. The relaxin-related factor may be RRF-1 or RRF-2.

The invention further provides for a relaxin-related factor ligandbody which

comprises a relaxin-related factor fused to an immunoglobulin constant region. The relaxin-related factor may be RRF-1 or RRF-2. In one embodiment of the relaxin-related factor ligandbody, the immunoglobulin constant region is the Fc portion of human IgG1. The invention further provides for a method of treatment, comprising administering a relaxin-related factor ligandbody to a human or animal body.

As used herein, the term "relaxin-related factor" includes the RRF-1 and RRF-2 molecules described herein, as well as functionally equivalent molecules in which amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Also included within the scope of the term "relaxin-related factor" are fragments or derivatives thereof which exhibit the same or similar biological activity and derivatives which are differentially modified during or after translation, e.g., by glycosylation, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc.

One skilled in the art will also recognize that the present invention provides DNA and RNA sequences that hybridize to a relaxin-related factor encoding sequence, under conditions of moderate stringency, as defined in, for example, Sambrook, et al. Molecular Cloning: A Laboratory Manual, 2 ed. Vol. 1, pp. 101-104, Cold Spring Harbor Laboratory Press (1989). Thus, a nucleic acid molecule contemplated by the invention includes one having a nucleotide sequence

deduced from an amino acid sequence of a relaxin-related factor as described herein, as well as a molecule having a sequence of nucleotides that hybridizes to such a nucleic acid under conditions of moderate stringency, and also a nucleotide sequence which is degenerate of the above sequences as a result of the genetic code, but which nonetheless encodes a relaxin-related factor.

A nucleic acid encoding a relaxin-related factor as described herein will facilitate the genetic engineering of cells to produce the relaxin-related factor by, e.g., transfection, transduction, electroporation, or microinjection of a suitable host cell using a nucleotide sequence encoding the relaxin-related factor in a suitable expression vector, thus forming a host cell vector system for the production of a polypeptide having the biological activity of the relaxin-related factor. A suitable host cell may be selected from the group consisting of bacterial cells (such as E. coli), yeast cells, fungal cells, insect cells and animal cells. Suitable animal cells include, but are not limited to, Vero cells, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells.

Any of the methods known to one skilled in the art for the insertion of DNA fragments into a vector may be used to construct expression vectors encoding the relaxin-related factor using appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include in vitro recombinant DNA and synthetic techniques and in vivo recombinations (genetic recombination). Expression of a nucleotide sequence encoding relaxin-related factor or peptide fragments thereof may be regulated by a second nucleotide sequence so that the protein or peptide is expressed in a host transformed with the recombinant DNA molecule. For example, expression of the relaxin-related factor described herein may be controlled by any promoter/enhancer element known in the art. Promoters which may be used to control expression of the relaxin-related factor include, but are not limited to the long terminal repeat as described in Squinto et al., (Cell 65:1-20 (1991)); the SV40 early promoter region (Bernoist and Chambon, Nature 290:304-310), the CMV promoter, the M-MuLV 5' terminal repeat, the promoter contained in the 3' long terminal repeat of Rous

sarcoma virus (Yamamoto, et al., Cell 22:787-797 (1980)), the herpes thymidine kinase promoter (Wagner et al., Proc. Natl. Acad. Sci. U.S.A. 78:144-1445 (1981)), the adenovirus promoter, the regulatory sequences of the metallothioein gene (Brinster et al., Nature 296:39-42 (1982)); prokaryotic expression vectors such as the β-lactamase promoter (Villa-Kamaroff, et al., Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731 (1978)), or the tac promoter (DeBoer, et al., Proc. Natl. Acad. Sci. U.S.A. 80:21-25 (1983)), see also "Useful proteins from recombinant bacteria" in Scientific American, 242:74-94 (1980); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADH (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., Cell 38:639-646 (1984); Ornitz et al., Cold Spring Harbor Symp. Quant. Biol. 50:399-409 (1986); MacDonald, Hepatology 7:425-515 (1987); insulin gene control region which is active in pancreatic beta cells (Hanahan, Nature 315:115-122 (1985)), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adames et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, Cell 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58); alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al, 1987, Genes and Devel. 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94); myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Shani, 1985, Nature 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378). The invention further provides for the production of antisense

compounds which are capable of specifically hybridizing with a sequence of RNA encoding the relaxin-related factor to modulate its expression. (Ecker, U.S. Patent No. 5,166,195, issued November 24, 1992).

Thus, according to the invention, expression vectors capable of being replicated in a bacterial or eukaryotic host comprising a nucleic acid encoding a relaxin-related factor are used to transfect a host and thereby direct expression of such nucleic acid to produce the relaxin-related factor, which may then be recovered in a biologically active form. As used herein, a biologically active form includes a form capable of binding to a receptor for the relaxin-related factor and causing a biological response such as a differentiated function or influencing the phenotype of the cell expressing the receptor.

Expression vectors containing the gene inserts can be identified by four general approaches: (a) DNA-DNA hybridization, (b) presence or absence of "marker" gene functions, (c) expression of inserted sequences and (d) PCR detection. In the first approach, the presence of a foreign gene inserted in an expression vector can be detected by DNA-DNA hybridization using probes comprising sequences that are homologous to an inserted relaxin-related factor encoding gene. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" gene functions (e.g., thymidine kinase activity, resistance to antibiotics, transformation phenotype, occlusion body formation in baculovirus, etc.) caused by the insertion of foreign genes in the vector. For example, if a nucleic acid encoding relaxin-related factor is inserted within the marker gene sequence of the vector, recombinants containing the insert can be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying the foreign gene product expressed by the recombinant. Such assays can be based, for example, on the physical or functional properties of a relaxin-related factor gene product, for example, by binding of the relaxin-related factor to its receptor or portion thereof which may be tagged with, for example, a detectable antibody or portion thereof or by binding to antibodies

produced against the relaxin-related factor or a portion thereof. Cells of the present invention may transiently, or preferably constitutively and permanently, express relaxin-related factor as described herein. In the fourth approach, DNA nucleotide primers can be prepared corresponding to a relaxin-related factor specific DNA sequence. These primers could then be used to PCR a relaxin-related factor gene fragment. (PCR Protocols: A Guide To Methods and Applications, Edited by Michael A. Innis et al., Academic Press (1990)).

A recombinant relaxin-related factor may be purified by any technique which allows for the subsequent formation of a stable, biologically active protein. For example, and not by way of limitation, a relaxin-related factor may be recovered from cells either as soluble proteins or as inclusion bodies, from which they may be extracted quantitatively by 8M guanidinium hydrochloride and dialysis. In order to further purify the relaxin-related factor, conventional ion exchange chromatography, hydrophobic interaction chromatography, reverse phase chromatography or gel filtration may be used.

In additional embodiments of the invention, a recombinant relaxin-related factor encoding gene may be used to inactivate or "knock out" the endogenous gene by homologous recombination, and thereby create a relaxin-related factor deficient cell, tissue, or animal. For example, and not by way of limitation, the recombinant relaxin-related factor encoding gene may be engineered to contain an insertional mutation, for example the neo gene, which would inactivate the native relaxin-related factor encoding gene. Such a construct, under the control of a suitable promoter, may be introduced into a cell, such as an embryonic stem cell, by a technique such as transfection, transduction, or injection. Cells containing the construct may then be selected by G418 resistance. Cells which lack an intact relaxin-related factor encoding gene may then be identified, e.g. by Southern blotting, PCR detection, Northern blotting or assay of expression. Cells lacking an intact relaxin-related factor encoding gene may then be fused to early embryo cells to generate transgenic animals deficient in such relaxin-related factor. Such an animal may be used to define specific in vivo processes,

normally dependent upon the relaxin-related factor.

The present invention also provides for antibodies to a relaxin-related factor described herein which are useful for detection of the factor in, for example, diagnostic applications. For the preparation of monoclonal antibodies directed toward a relaxin-related factor, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in "Monoclonal Antibodies and Cancer Therapy," Alan R. Liss, Inc. pp. 77-96) and the like are within the scope of the present invention.

Monoclonal antibodies for diagnostic or therapeutic use may be human monoclonal antibodies or chimeric human-mouse (or other species) monoclonal antibodies. Human monoclonal antibodies may be made by any of numerous techniques known in the art (e.g., Teng et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:7308-7312; Kozbor et al., 1983, Immunology Today 4:72-79; Olsson et al., 1982, Meth. Enzymol. 92:3-16). Chimeric antibody molecules may be prepared containing a mouse antigen-binding domain with human constant regions (Morrison et al., 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851, Takeda et al., 1985, Nature 314:452).

Various procedures known in the art may be used for the production of polyclonal antibodies to epitopes of a relaxin-related factor described herein. For the production of antibody, various host animals can be immunized by injection with a relaxin-related factor, or a peptide fragment or derivative thereof, including but not limited to rabbits, mice and rats. Various adjuvants may be used to increase the immunological response, depending on the host species and including, but not limited to, Freund's adjuvant (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as

lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (Bacille Calmette-Guerin) and Corynebacterium parvum.

A molecular clone of an antibody to a selected relaxin-related factor epitope can be prepared by known techniques. Recombinant DNA methodology (see e.g., Maniatis et al., 1982, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) may be used to construct nucleotide sequences which encode a monoclonal antibody molecule, or antigen binding region thereof.

The present invention provides for antibody molecules as well as fragments of such antibody molecules. Antibody fragments which contain the idiotype of the molecule can be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragment; and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent. Antibody molecules may be purified by known techniques, e.g., immunoabsorption or immunoaffinity chromatography, chromatographic methods such as HPLC (high performance liquid chromatography), or a combination thereof.

Having isolated a nucleic acid molecule encoding a relaxin-related factor, it should pose no great difficulty to create an expression construct that would yield a secreted protein consisting of the entire coding sequence of the relaxin-related factor fused to the human immunoglobulin gamma-1 constant region (IgG1 Fc). These fusion proteins are called "relaxin-related factor ligandbodies." The relaxin-related factor may be RRF-1 or RRF-2.

The Fc portion may be prepared as follows. A DNA fragment encoding the Fc portion of human IgG1 that spans from the hinge region to the carboxy-terminus

of the protein may be amplified from human placental cDNA by PCR with oligonucleotides corresponding to the published sequence of human IgG1; the resulting DNA fragment may be cloned into a plasmid vector. Appropriate DNA restriction fragments from a plasmid encoding the full-length relaxin-related factor and from the human IgG1 Fc plasmid may be ligated on either side of a short PCR-derived fragment that is designed so as to fuse, in-frame, the chosen relaxin-related factor with human IgG1 Fc protein-coding sequences.

The present invention therefore also provides for a relaxin-related factor ligandbody which specifically binds the chosen relaxin-related factor's receptor. In one embodiment, the present invention provides for a relaxin-related factor ligandbody which comprises a relaxin-related factor ligand fused to an immunoglobulin constant region. In a further embodiment, the ligandbody comprises a relaxin-related factor and the Fc portion of human IgG1. In a further embodiment, the ligandbody comprises RRF-1 and the Fc portion of human IgG1. In yet another embodiment, the ligandbody comprises RRF-2 and the Fc portion of human IgG1. A relaxin-related factor ligandbody as described above may be used in a method of treatment of the human or animal body, or in a method of diagnosis.

By way of nonlimiting example, relaxin-related factor ligandbodies may be useful for the delivery of toxins to a receptor bearing cell. Alternatively, other molecules, such as growth factors, cytokines or nutrients, may be delivered to a relaxin-related factor receptor bearing cell via relaxin-related factor ligandbodies. A relaxin-related factor ligandbody could also be used as a diagnostic reagent for a relaxin-related factor receptor, to detect a receptor in vivo or in vitro. A relaxin-related factor ligandbody may be useful as a diagnostic reagent for detecting a relaxin-related factor receptor by, for example, tissue staining or whole body imaging.

The present invention also provides for pharmaceutical compositions comprising a relaxin-related factor ligand or ligandbody described herein, peptide

fragments thereof, or derivatives in a pharmacologically acceptable vehicle. The relaxin-related factor protein, or peptide fragments or derivatives thereof, may be administered systemically or locally. Any appropriate mode of administration known in the art may be used including, but not limited to, intravenous, intrathecal, intraarterial, intranasal, oral, subcutaneous, intraperitoneal, or by local injection or surgical implant. Sustained release formulations are also provided for.

This invention will be better understood from the Examples which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are not limiting but are merely illustrative of the invention.

EXAMPLE 1 - CLONING OF A MOLECULE RELATED TO RELAXIN AND THE INSULIN FAMILY OF LIGANDS

A search of the database of expressed sequence tags using the sequence of relaxin (a member of the insulin family, which also includes insulin, IGF1, IGF2, and "Leydig insulin-like protein") as the test sequence revealed the existence of a sequence, found in human testis, related to relaxin at the amino acid level. The hallmarks of the insulin family of proteins consist mainly of a particular arrangement of cysteine residues, two near the amino terminus and four towards the carboxy terminus, which form a characteristic structure of disulfide bonds. The sequence found on the database contained the first two cysteines, along with a few other residues in their immediate vicinity that are also characteristic of the insulin family. This homology was rather weak, and the presence of a stop codon in the middle of the database sequence, with no evidence for the characteristic C-terminal cysteines, raised further doubts about whether this sequence represented an authentic new insulin family member.

Oligonucleotides were synthesized corresponding to the database sequence and

were assembled by PCR to constitute a 400-base probe. A human testis cDNA library was constructed and then screened with this synthetic probe. Several clones resulting from this screen were sequenced, and it was found that at least one of the clones contained an open reading frame whose predicted protein sequence displayed all the expected features of a new insulin family member, particularly with regard to the cluster of four cysteines at the C-terminus of the molecule. The nucleotide and deduced amino acid (single letter code) sequences of the molecule are shown in Figure 1. Overall, the molecule is more closely related to relaxin than to any of the other insulin family members. The molecule was therefore called relaxin-related factor-1 (RRF-1). Figure 2 shows a comparative alignment of the relaxin-related factor-1 amino acid sequence with the sequences of other members of the insulin family.

Northern blot analysis was performed by running samples of adult human poly A RNA on a 1% agarose-formaldehyde gel, transferring to a nylon membrane and probing with a 400 base pair probe derived from the nucleotide sequence of the relaxin-related factor. As shown in Figure 3, the message for RRF-1 was highly expressed in human testis, but was not detected in any of the other tissue samples that were assayed, including brain, cerebellum, thymus, heart, liver, small intestine, skeletal muscle, bone marrow, prostate, ovary, uterus, and placenta. Because production of RRF-1 appears to be specific to testis, it may be involved in the maturation of sperm and therefore may have a role in treatment of fertility disorders.

EXAMPLE 2 - EXPRESSION OF HUMAN RELAXIN-RELATED FACTOR-1 cDNA

Expression vectors were constructed containing the human relaxin-related factor-1 cDNA. One version had a triple-myc tag inserted at its C-terminus to make the protein detectable. COS-7 cells were transiently transfected with either the expression vector or a control vector by the DEAE-dextran transfection protocol.

Briefly, COS-7 cells were plated at a density of 1.0 x 10^6 cells per 100 mm plate 24 hours prior to transfection. For transfection, the cells were cultured in serum-free DMEM containing 400 μg/ml of DEAE-dextran, 1 μM chloroquine, and 2 mM glutamine, and 1 μg of the appropriate DNA for 3-4 hours at 37°C in an atmosphere of 5% CO_2. The transfection media was aspirated and replaced with phosphate-buffered saline with 10% DMSO for 2-3 min. Following this DMSO "shock", the COS-7 cells were placed into DMEM with 1% each of penicillin and streptomycin, and 2 mM glutamine. COS media containing secreted ligand was harvested after three days. Expression of the myc-tagged version was verified by Western blot analysis of the supernatants, using antibodies against the myc tag as a probe.

EXAMPLE 3 - CLONING OF A SECOND MOLECULE RELATED TO THE INSULIN FAMILY OF LIGANDS

A second search of the database of expressed sequence tags using the sequence of relaxin as the test sequence revealed the existence of numerous sequence entries, all found in human placental cDNA and representing essentially the same sequence, which was related to relaxin at the amino acid level. As stated previously, the hallmarks of the insulin family of proteins consist mainly of a particular arrangment of cysteine residues, two near the amino terminus and four towards the carboxy terminus, which form a characteristic structure of disulfide bonds. The human placental sequences found on the database contained the last four cysteines, along with a few other residues in their immediate vicinity that are also characteristic of the insulin family. This homology was rather weak, and there was variation among the different entries on the database, with some having frameshift mutations with respect to others, making it somewhat unclear whether the homology represented an authentic new insulin family member. In addition, it was noticed that many of these sequences had a stretch that was identical to the cDNA sequence for a known ribosomal protein, L22, and these sequence entries were identified on the database as L22, further obscuring the homology with relaxin and supporting the

possibility that the relaxin homology was artifactual. Further examination revealed that the L22 protein was transcribed in the opposite direction from the relaxin-homologous sequence, so it was possible that there were two overlapping genes transcribed from opposite strands, and splicing mechanisms may have included portions of one cDNA in the other.

Oligonucleotides were synthesized corresponding to parts of the database sequence and were used to screen a human placental cDNA library. One clone resulting from this screen was sequenced, and it was found that it contained an open reading frame whose predicted protein sequence displayed many of the expected features of a new insulin family member: there was a signal sequence at the amino terminus, allowing for the secretion of the protein, followed by a sequence with two cysteines separated by 11 amino acid residues, as is the case with all other insulin family members. In addition, there was a tetrabasic sequence just preceding the C-terminal four-cysteine cluster, which is present in relaxin and is probably involved in proteolytic processing of the protein (other members of the family have dibasic residues at this position). Although these sequence signatures identify this new molecule as an insulin family member, the sequence in the vicinity of the amino-terminal cysteine cluster is rather more distant from all the others, and therefore the new molecule may be thought of as identifying a new branch of the insulin family. Overall, the new molecule is more closely related to relaxin than to any of the other insulin family members. The molecule was therefore called relaxin-related factor-2 (RRF-2). Figure 2 shows a comparative alignment of the relaxin-related factor-2 amino acid sequence with the sequences of RRF-1 and other members of the insulin family.

Northern blot analysis was performed as previously described in Example 1 and the results, shown in Figure 5, revealed that the message for relaxin-related factor-2 (RRF-2) was highly expressed in human placenta, but was not detected in any of the other samples that were assayed. Because production of RRF-2 appears to be specific to placenta, it may be useful for modulating the reproductive physiology of mammals during pregnancy and parturition.

EXAMPLE 4 - EXPRESSION OF HUMAN RELAXIN-RELATED FACTOR-2 cDNA

Expression vectors were constructed containing the human relaxin-related factor-2 cDNA. One version had a FLAG tag inserted at its N-terminus to make the protein detectable. COS-7 cells were transiently transfected with either the expression vector or a control vector by the DEAE-dextran transfection protocol. Expression of the FLAG-tagged version was verified by staining of the cells using antibodies against the FLAG tag as a probe.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying Figures. Such modifications are intended to fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule encoding a relaxin-related factor.

2. An isolated nucleic acid molecule according to claim 1, having a sequence selected from the group consisting of:
   (a) the nucleotide sequence comprising the coding region of the relaxin-related factor as set forth in Figure 1;
   (b) a nucleotide sequence that hybridizes under moderately stringent conditions to the nucleotide sequence of (a) and which encodes a molecule having the biological activity of the relaxin-related factor; or
   (c) a nucleotide sequence which, but for the degeneracy of the genetic code would hybridize to a nucleotide sequence of (a) or (b), and which encodes a molecule having the biological activity of the relaxin-related factor.

3. A vector which comprises a nucleic acid molecule of claim 1 or 2.

4. A vector according to claim 3, wherein the nucleic acid molecule is operatively linked to an expression control sequence capable of directing its expression in a host cell.

5. A plasmid according to claim 3 or 4.

6. Isolated relaxin-related factor substantially free of other proteins.

7. Isolated relaxin-related factor-1 substantially free of other proteins, encoded by a nucleic acid molecule according to claim 2.

8. A host-vector system for the production of a relaxin-related factor which comprises a vector of claim 3 or 4, in a host cell.

9. A host-vector system according to claim 8, wherein the host cell is a bacterial, yeast, insect or mammalian cell.

10. A method of producing a relaxin-related factor which comprises growing cells of a host-vector system of claim 8 or 9, under conditions permitting production of the relaxin-related factor, and recovering the relaxin-related factor so produced.

11. An antibody which specifically binds the relaxin-related factor of claim 6 or 7.

12. An antibody according to claim 11, which is a monoclonal antibody.

13. A pharmaceutical composition comprising a relaxin-related factor according to claim 6 or 7, and a pharmaceutically acceptable carrier.

14. A pharmaceutical composition comprising an antibody according to claim 11 or 12, and a pharmaceutically acceptable carrier.

15. A relaxin-related factor according to claim 6 or 7, an antibody according to claim 11 or 12, or a composition according to claim 13 or 14, for use in a method of treatment of the human or animal body, or in a method of diagnosis.

16. A polypeptide produced by the method of claim 10.

17. A ligandbody which comprises relaxin-related factor-1 fused to an immunoglobulin constant region.

18. The ligandbody of claim 17, wherein the immunoglobulin constant region is the Fc portion of human IgG1.

19. A ligandbody according to claim 17 or 18, for use in a method of treatment of the human or animal body, or in a method of diagnosis.

20. An isolated nucleic acid molecule according to claim 1, having a sequence selected from the group consisting of:
    (a) the nucleotide sequence comprising the coding region of the relaxin-related factor as set forth in Figure 4;
    (b) a nucleotide sequence that hybridizes under moderately stringent conditions to the nucleotide sequence of (a) and which encodes a molecule having the biological activity of the relaxin-related factor; or
    (c) a nucleotide sequence which, but for the degeneracy of the genetic code would hybridize to a nucleotide sequence of (a) or (b), and which encodes a molecule having the biological activity of the relaxin-related factor.

21. A vector which comprises a nucleic acid molecule of claim 20.

22. A vector according to claim 21, wherein the nucleic acid molecule is operatively linked to an expression control sequence capable of directing its expression in a host cell.

23. A plasmid according to claim 21 or 22.

24. Isolated relaxin-related factor-2 substantially free of other proteins, encoded by a nucleic acid molecule according to claim 20.

25. A host-vector system for the production of a relaxin-related factor which comprises a vector of claim 21 or 22, in a host cell.

26. A host-vector system according to claim 25, wherein the host cell is a

bacterial, yeast, insect or mammalian cell.

27. A method of producing a relaxin-related factor which comprises growing cells of a host-vector system of claim 25 or 26, under conditions permitting production of the relaxin-related factor, and recovering the relaxin-related factor so produced.

28. An antibody which specifically binds the relaxin-related factor of claim 24.

29. An antibody according to claim 28, which is a monoclonal antibody.

30. A pharmaceutical composition comprising a relaxin-related factor according to claim 24, and a pharmaceutically acceptable carrier.

31. A pharmaceutical composition comprising an antibody according to claim 28 or 29, and a pharmaceutically acceptable carrier.

32. A relaxin-related factor according to claim 24, an antibody according to claim 28 or 29, or a composition according to claim 30 or 31, for use in a method of treatment of the human or animal body, or in a method of diagnosis.

33. A polypeptide produced by the method of claim 27.

34. A ligandbody which comprises relaxin-related factor-2 fused to an immunoglobulin constant region.

35. The ligandbody of claim 34, wherein the immunoglobulin constant region is the Fc portion of human IgG1.

36. A ligandbody according to claim 34 or 35, for use in a method of treatment of the human or animal body, or in a method of diagnosis.

Fig.1.

```
          10         20         30         40         50         60
     *    *     *    *     *    *     *    *     *    *     *    *
CCATTGTGCT GGAAAGCTCG GCCTGGCGCA GGCGCAGACA GGGAGCAGGG TCCGGCCGAC
GGTAACACGA CCTTTCGAGC CGGACCGCGT CCGCGTCTGT CCCTCGTCCC AGGCCGGCTG

          70         80         90        100         110
     *    *     *    *     *    *     *    *       *    *      *
CGCCATTGCA CAACGCGGAG GACTAGCCTG GGGTCACAGG G ATG CCG CGG CTC CTC
GCGGTAACGT GTTGCGCCTC CTGATCGGAC CCCAGTGTCC C TAC GGC GCC GAG GAG
                                                M   P   R   L   L>

  120       130       140       150       160
   *    *    *    *    *    *    *    *    *
CGC TTG TCC CTG CTG TGG CTT GGA CTC CTG CTG GTT CGG TTT TCT CGT
GCG AAC AGG GAC GAC ACC GAA CCT GAG GAC GAC CAA GCC AAA AGA GCA
 R   L   S   L   L   W   L   G   L   L   L   V   R   F   S   R>

      170       180       190       200       210
 *    *    *    *    *    *    *    *    *    *
GAA CTG AGC GAC ATC AGC AGT GCC AGG AAG CTG TGC GGC AGG TAC TTG
CTT GAC TCG CTG TAG TCG TCA CGG TCC TTC GAC ACG CCG TCC ATG AAC
 E   L   S   D   I   S   S   A   R   K   L   C   G   R   Y   L>

        220       230       240       250       260
  *      *    *    *    *    *    *    *    *    *
GTG AAA GAA ATA GAA AAA CTC TGC GGC CAT GCC AAC TGG AGC CAG TTC
CAC TTT CTT TAT CTT TTT GAG ACG CCG GTA CGG TTG ACC TCG GTC AAG
 V   K   E   I   E   K   L   C   G   H   A   N   W   S   Q   F>
```

Fig.1 (Cont 1/3).

```
                270         280         290         300
         *         *     *     *     *     *     *     *     *
CGT TTC GAG GAG GAA ACC CCT TTC TCA CGG TTG ATT GCA CAG GCC TCG
GCA AAG CTC CTC CTT TGG GGA AAG AGT GCC AAC TAA CGT GTC CGG AGC
 R   F   E   E   E   T   P   F   S   R   L   I   A   Q   A   S>

310         320         330         340         350
 *     *     *     *     *     *     *     *     *     *
GAG AAG GTC GAA GCC TAC AGC CCA TAC CAG TTC GAA AGC CCG CAA ACC
CTC TTC CAG CTT CGG ATG TCG GGT ATG GTC AAG CTT TCG GGC GTT TGG
 E   K   V   E   A   Y   S   P   Y   Q   F   E   S   P   Q   T>

    360         370         380         390         400
     *     *     *     *     *     *     *     *     *
GCT TCC CCG GCC CGG GGA AGA GGC ACA AAC CCA GTG TCT ACT TCT TGG
CGA AGG GGC CGG GCC CCT TCT CCG TGT TTG GGT CAC AGA TGA AGA ACC
 A   S   P   A   R   G   R   G   T   N   P   V   S   T   S   W>

       410         420         430         440         450
 *     *     *     *     *     *     *     *     *     *
GAA GAA GCA GTA AAC AGT TGG GAA ATG CAG TCA CTA CCT GAG TAT AAG
CTT CTT CGT CAT TTG TCA ACC CTT TAC GTC AGT GAT GGA CTC ATA TTC
 E   E   A   V   N   S   W   E   M   Q   S   L   P   E   Y   K>

          460         470         480         490         500
     *     *     *     *     *     *     *     *     *     *
GAT AAA AAG GGA TAT TCA CCC CTT GGT AAG ACA AGA GAA TTT TCT TCA
CTA TTT TTC CCT ATA AGT GGG GAA CCA TTC TGT TCT CTT AAA AGA AGT
 D   K   K   G   Y   S   P   L   G   K   T   R   E   F   S   S>
```

Fig.1 (Cont 2/3).

```
             510         520          530         540
     *      *     *      *      *     *      *     *     *
TCA CAT AAT ATC AAT GTA TAT ATT CAT GAG AAT GCA AAA TTT CAG AAG
AGT GTA TTA TAG TTA CAT ATA TAA GTA CTC TTA CGT TTT AAA GTC TTC
 S   H   N   I   N   V   Y   I   H   E   N   A   K   F   Q   K>

550         560         570         580         590
 *      *    *      *      *     *     *      *     *       *
AAA CGT AGA AAC AAA ATT AAA ACC TTA AGC AAT TTG TTT TGG GGG CAT
TTT GCA TCT TTG TTT TAA TTT TGG AAT TCG TTA AAC AAA ACC CCC GTA
 K   R   R   N   K   I   K   T   L   S   N   L   F   W   G   H>

   600         610         620         630         640
     *     *     *     *     *      *     *     *     *
CAT CCC CAA AGA AAA CGC AGA GGA TAT TCA GAA AAG TGT TGT CTT ACA
GTA GGG GTT TCT TTT GCG TCT CCT ATA AGT CTT TTC ACA ACA GAA TGT
 H   P   Q   R   K   R   R   G   Y   S   E   K   C   C   L   T>

     650         660         670         680         690
 *    *      *     *     *     *      *     *     *     *
GGA TGT ACA AAA GAA GAA CTT AGC ATT GCA TGT CTT CCA TAT ATT GAT
CCT ACA TGT TTT CTT CTT GAA TCG TAA CGT ACA GAA GGT ATA TAA CTA
 G   C   T   K   E   E   L   S   I   A   C   L   P   Y   I   D>
```

Fig.1 (Cont 3/3).

```
        700            710             720           730            740
 *       *       *      *       *       *      *       *       *      *
TTT AAA AGG CTA AAG GAA AAA AGA TCA TCA CTT GTA ACT AAG ATA TAC
AAA TTT TCC GAT TTC CTT TTT TCT AGT AGT GAA CAT TGA TTC TAT ATG
 F   K   R   L   K   E   K   R   S   S   L   V   T   K   I   Y>

        750          760          770         780         790
   *     *      *     *      *     *     *     *     *     *     *
TAACCATCTT AGAATTTTTT CTAACCTAAT AAAAGCTTTC CAGCTTTCCA GCACAATGG
ATTGGTAGAA TCTTAAAAAA GATTGGATTA TTTTCGAAAG GTCGAAAGGT CGTGTTACC
```

Fig.2.

```
                                         *               *                                   **   *        *
?-1      M  (SIGNAL SEQUENCE) ------RKLCGRYLVKEIEKLCG  (123aa) RKRR ----GYSEK CCLTGCTKEELSIACLPYIDEKRLKEKRSSLVTKIY
?-2      M  (SIGNAL SEQUENCE) -----ELRG..PRFGKHLLSY.P  ( 66aa) RKKR SGRHRFDPF ..EVI.DDGTSVKL.T

AXINH1   .  (SIGNAL SEQUENCE) KWKDDVI.....E..RAQIAI..  (110aa) KKRR P-YVALF.- ...I....RS.AKY.
AXINH2   .  (SIGNAL SEQUENCE) -WMEEVI.....E..RAQIAI..  (109aa) RKKR QLYSALAN. ..HV....RS.ARF.
(DIG     .  (SIGNAL SEQUENCE) PTPEMRE.....HHF.RALVRV..  ( 54aa) RHHR AAATNPARY ...S...QQD.LTL.-..
SULIN    .  (SIGNAL SEQUENCE) ---FVNQH...SH..EALYLV..  ( 43aa) KR-- ----GIVEQ ..TSI.SLYQ.ENY.
?1       .  (SIGNAL SEQUENCE) --TAGPET...AE..DALQFV..  ( 16aa) RR-- APQTGIVDE ..FRS.DLRR.EMY.  (86aa)
?2       .  (SIGNAL SEQUENCE) --YRPSET...GE..DTLQFV..  ( 14aa) RR-- SR--GIVEE ..FRS.DLAL.ETY.  (96aa)
```

Fig 3.

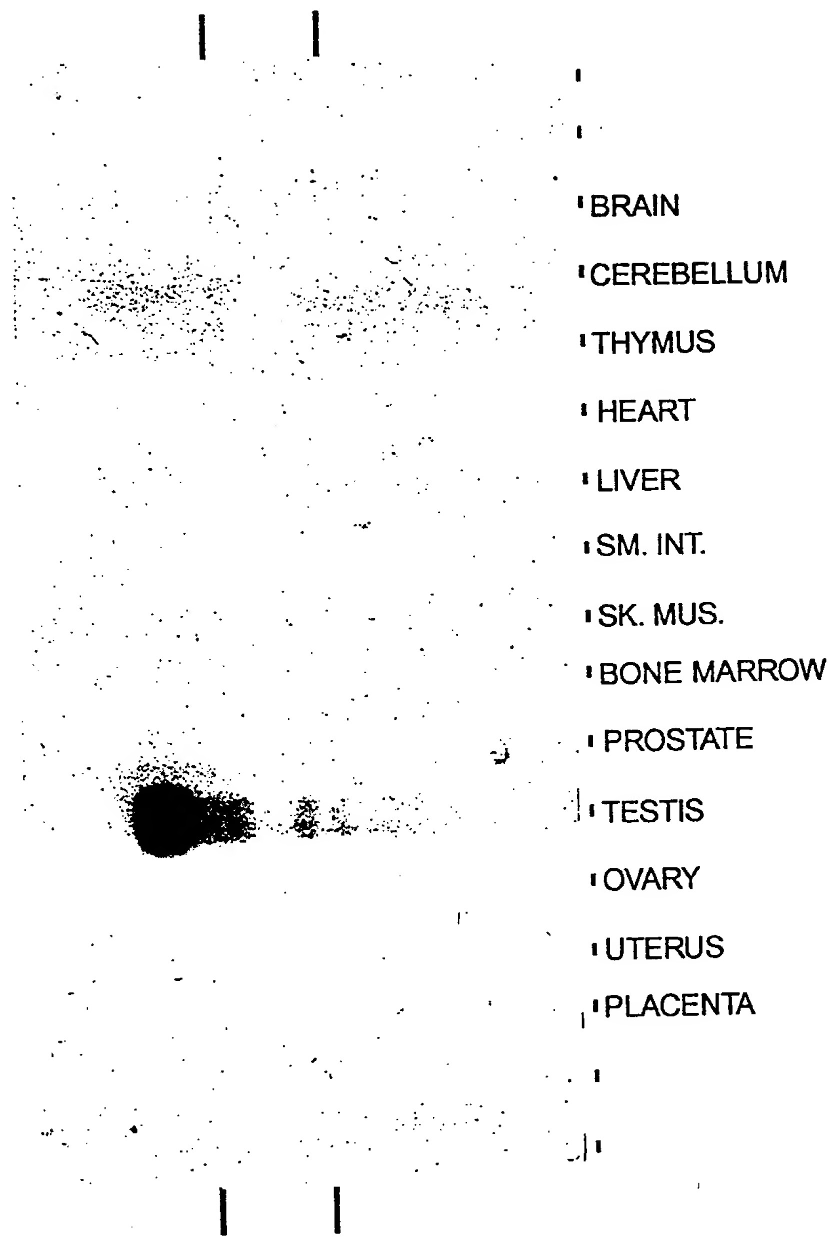

BRAIN
CEREBELLUM
THYMUS
HEART
LIVER
SM. INT.
SK. MUS.
BONE MARROW
PROSTATE
TESTIS
OVARY
UTERUS
PLACENTA

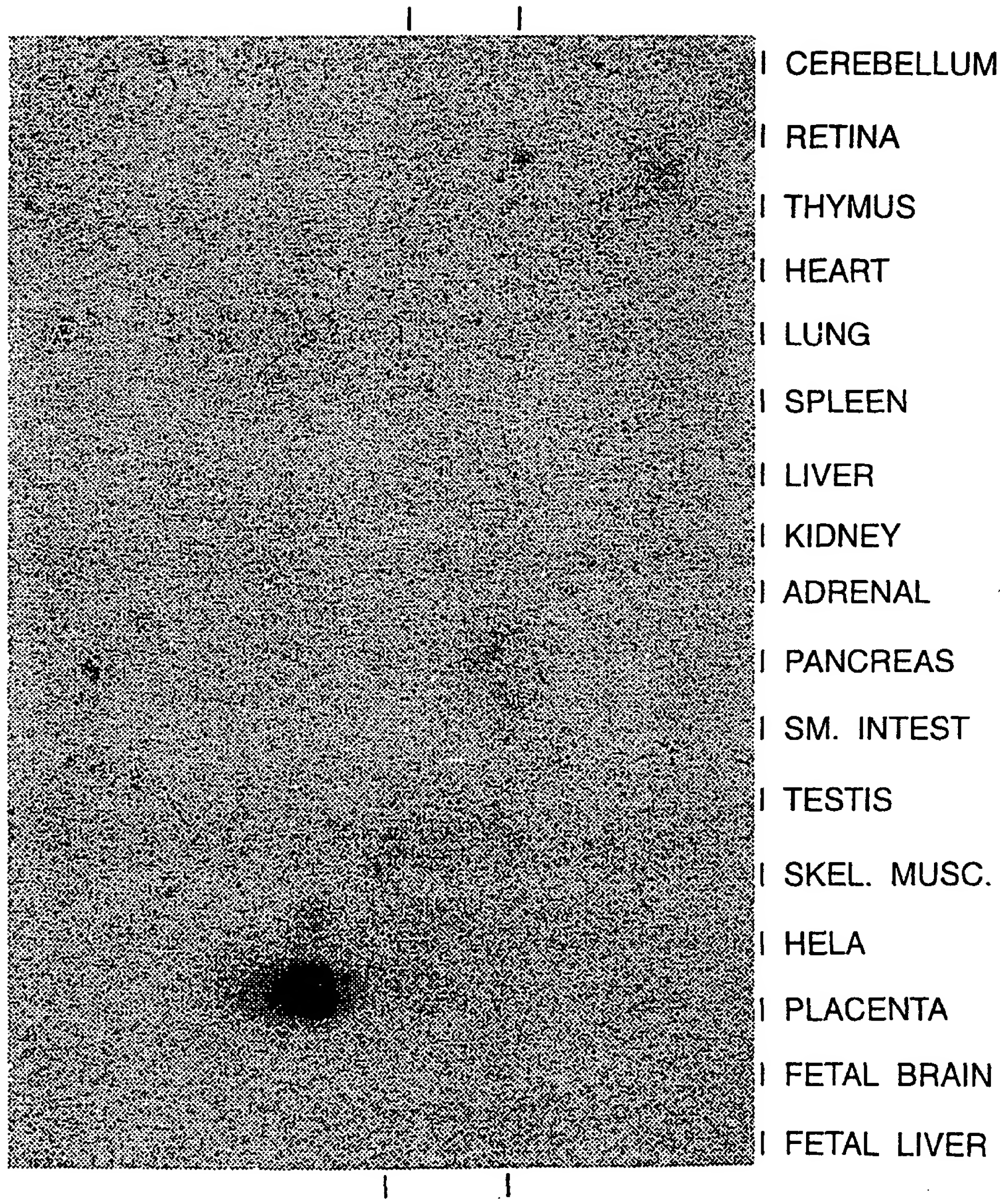
Fig.4.

```
        10         20         30         40         50         60         70         80
         *          *          *          *          *          *          *          *
CCATTGTGCT GGAAAGGACA CACCAGCACA GTCTGGTAGG CTACAGCAGC AAGTCTCTAA AGAAAGGCTG AGAACACCCA
GGTAACACGA CCTTTCCTGT GTGGTCGTGT CAGACCATCC GATGTCGTCG TTCAGAGATT TCTTTCCGAC TCTTGTGGGT

        90        100        110        120        130        140
         *          *          *          *          *          *
GAACAGGAGA GTTCAGGTCC AGG ATG GCC AGC CTG TTC CGG TCC TAT CTG CCA GCA ATC TGG CTG CTG
CTTGTCCTCT CAAGTCCAGG TCC TAC CGG TCG GAC AAG GCC AGG ATA GAC GGT CGT TAG ACC GAC GAC
                          M   A   S   L   F   R   S   Y   L   P   A   I   W   L   L>

150        160        170        180        190        200        210
  *          *          *          *          *          *          *
CTG AGC CAA CTC CTT AGA GAA AGC CTA GCA GCA GAG CTG AGG GGA TGT GGT CCC CGA TTT GGA AAA
GAC TCG GTT GAG GAA TCT CTT TCG GAT CGT CGT CTC GAC TCC CCT ACA CCA GGG GCT AAA CCT TTT
 L   S   Q   L   L   R   E   S   L   A   A   E   L   R   G   C   G   P   R   F   G   K>

    220        230        240        250        260        270        280
      *          *          *          *          *          *          *
CAC TTG CTG TCA TAT TGC CCC ATG CCT GAG AAG ACA TTC ACC ACC ACC CCA GGA GGG TGG CTG CTG
GTG AAC GAC AGT ATA ACG GGG TAC GGA CTC TTC TGT AAG TGG TGG TGG GGT CCT CCC ACC GAC GAC
 H   L   L   S   Y   C   P   M   P   E   K   T   F   T   T   T   P   G   G   W   L   L>

       290        300        310        320        330        340
         *          *          *          *          *          *
GAA TCT GGA CGT CCC AAA GAA ATG GTG TCA ACC TCC AAC AAC AAA GAT GGA CAA GCC TTA GGT ACG
CTT AGA CCT GCA GGG TTT CTT TAC CAC AGT TGG AGG TTG TTG TTT CTA CCT GTT CGG AAT CCA TGC
 E   S   G   R   P   K   E   M   V   S   T   S   N   N   K   D   G   Q   A   L   G   T>
```

Fig.4 (Cont).

```
 350         360         370          380         390         400          410
   *    *     *     *     *     *     *     *     *     *     *     *     *
ACA TCA GAA TTC ATT CCT AAT TTG TCA CCA GAG CTG AAG AAA CCA CTG TCT GAA GGG CAG CCA TCA
TGT AGT CTT AAG TAA GGA TTA AAC AGT GGT CTC GAC TTC TTT GGT GAC AGA CTT CCC GTC GGT AGT
 T   S   E   F   I   P   N   L   S   P   E   L   K   K   P   L   S   E   G   Q   P   S>

       420         430         440         450         460         470
 *      *     *     *     *     *     *     *     *     *     *     *     *
TTG AAG AAA ATA ATA CTT TCC CGC AAA AAG AGA AGT GGA CGT CAC AGA TTT GAT CCA TTC TGT TGT
AAC TTC TTT TAT TAT GAA AGG GCG TTT TTC TCT TCA CCT GCA GTG TCT AAA CTA GGT AAG ACA ACA
 L   K   K   I   I   L   S   R   K   K   R   S   G   R   H   R   F   D   P   F   C   C>

480         490          500         510         520        530        540        550
 *     *     *     *     *     *     *     *     *     *     *     *     *     *
GAA GTA ATT TGT GAC GAT GGA ACT TCA GTT AAA TTA TGT ACA TAGTAGAGTA ATCATGGACT GGACATCTCA
CTT CAT TAA ACA CTG CTA CCT TGA AGT CAA TTT AAT ACA TGT ATCATCTCAT TAGTACCTGA CCTGTAGAGT
 E   V   I   C   D   D   G   T   S   V   K   L   C   T>

       560        570        580        590        600        610        620        630
  *     *     *     *     *     *     *     *     *     *     *     *     *     *     *
TCCATTCTCA TATGTATTCT CAATGACAAA TTCACTGATG CCCAATTAAA TGATTGCTGT TTATTAGAAC ATGAAAAAAA
AGGTAAGAGT ATACATAAGA GTTACTGTTT AAGTGACTAC GGGTTAATTT ACTAACGACA AATAATCTTG TACTTTTTTT

       640
  *     *     *
AAACTTTCCA GCACAATGG
TTTGAAAGGT CGTGTTACC
```

Fig.5.


CEREBELLUM
RETINA
THYMUS
HEART
LUNG
SPLEEN
LIVER
KIDNEY
ADRENAL
PANCREAS
SM. INTEST
TESTIS
SKEL. MUSC.
HELA
PLACENTA
FETAL BRAIN
FETAL LIVER